EM/DT (Q11L12)

Total Questions: 20

Most Correct Answers: **#13**

Least Correct Answers: **#16**

1. **With the increase of k, the decision boundary will be**

4/8 (A) simplified

3/8 (B) more complex

1/8 (C) I do not know

0/8 (D) unchanged

2. **Which of these algorithms can be used to fill the missing values**

2/8 (A) KNN for regression

4/8 (B) KNN for classification

1/8 (C) both

1/8 (D) I do not know

3. **Decision Tree Decision Boundaries**

5/8 (A) are a step-wise constant function

0/8 (B) I do not know

0/8 (C) continuous function

3/8 (D) are axis-parallel rectangles

4. **Root Node has**

3/8 (A) no incoming edges and zero or more outgoing edges

3/8 (B) one incoming edge and two or more outgoing edges

1/8 (C) one incoming edge and no outgoing edges

1/8 (D) I do not know

5. **Pruning the tree means**

6/8 (A) Simplify the tree

1/8 (B) Split the tree's nodes

0/8 (C) Merge the tree's nodes

1/8 (D) I do not know

6. **Gini index equals to**

2/8 (A) 1 - sum (pi^2)

0/8 (B) 1 + sum (pi^2)

2/8 (C) sum(pi * log(pi))

0/8 (D) -sum(pi * log(pi))

4/8 (E) I do not know

7. **Entropy starts with 0**

5/8 (A) True

2/8 (B) False

1/8 (C) I do not know

8. **Overall impurity measure can be obtained by**

3/8 (A) a weighted average of individual rectangles

1/8 (B) majority voting

4/8 (C) I do not know

9. **At each stage, we choose the split with**

3/8 (A) the lowest Gini index

1/8 (B) the lowest Chi-square value

2/8 (C) the highest entropy

2/8 (D) I do not know

10. **We can perform the Decision Trees in r using**

4/8 (A) rpart()

0/8 (B) decisiontree()

2/8 (C) destree()

1/8 (D) reg.tree()

1/8 (E) I do not know

11. **minsplit in R means**

5/8 A the minimum number of observations that must exist in a node in order for a split to be attempted

0/8 B the minimum number of observations in any terminal node

2/8 C the minimum number of splits

1/8 D I do not know

12. **Bagging is a technique used to reduce**

2/8 A the variance of our predictions

1/8 B the bias of our predictions

3/8 C both

2/8 D I do not know

13. **Bootstrap aggregation allows sampling**

6/8 A with replacement

1/8 B without replacement

1/8 C I do not know

0/8 D both

14. **How can Ensemble methods be constructed?**

1/8 A By manipulating the training set

0/8 B By manipulating the input features

0/8 C By manipulating the class labels

0/8 D By manipulating the learning algorithm

1/8 E All of them

0/8 F None

6/8 G I do not know

15. **Repeatedly sampling observations are taken**

2/8 A from general population

4/8 B original sample data set

2/8 C I do not know

0/8 D None

16. Random Forest differs from bagging

0/8 A by a random sample of m predictors

5/8 B by bootstrapped training samples

1/8 C by adaptive sampling

2/8 D I do not know

17. Boosting differs from bagging

0/8 A by a random sample of m predictors

2/8 B by bootstrapped training samples

1/8 C by adaptive sampling

5/8 D I do not know

18. Averaging many highly correlated quantities

3/8 A lead to as large of a reduction in variance

1/8 B does not lead to as large of a reduction in variance

2/8 C lead to as large of a reduction in bias

2/8 D I do not know

19. We can perform a Random forest in R using the function

3/8 A randomForest()

0/8 B rf()

1/8 C randomF()

3/8 D boot()

1/8 E I do not know

20. Random Forest works

2/8 A for classification

2/8 B for regression

3/8 C both

1/8 D I do not know